In silico design of *Mycobacterium tuberculosis* epitope ensemble vaccines

Preksha Shah [1], Jaymisha Mistry [1],

Pedro A Reche [2],

Derek Gatherer [3],

& Darren R Flower [1,*]

[1] School of Life and Health Sciences, Aston University, Aston Triangle, Birmingham, United Kingdom, B4 7ET.

[2] Immunomedicine Group, Facultad de Medicina, Departamento de Microbiologia I, Universidad Complutense de Madrid, Madrid, Spain.

[3] Division of Biomedical & Life Sciences, Faculty of Health & Medicine, Lancaster University, Lancaster LA1 4YW, UK.

* Corresponding author

d.r.flower@aston.ac.uk

44 121 204 5182

Abstract

Effective control of *Mycobacterium tuberculosis* is a global necessity. In 2015, tuberculosis (TB) caused more deaths than HIV. Considering the increasing prevalence of multi-drug resistant forms of *M. tuberculosis*, the need for effective TB vaccines is imperative. Currently, the only licensed TB vaccine is *Bacillus Calmette-Guérin* (BCG). Yet, BCG has many drawbacks limiting its efficacy and applicability. We appl~~y~~ied advanced computational procedures to derive a universal TB vaccine and one targeting East Africa. Our approach selects an optimal set of highly conserved, experimentally validated epitopes, with high projected population coverage (PPC). Through rigorous data analysis, five different potential vaccine combinations were selected each with PPC above 80% for East Africa and above 90% for the World. Two potential vaccines only contained CD8+ epitopes ~~only~~, while the others included both~~,~~ CD4+ and CD8+ epitopes. Our prime vaccine candidate was a putative seven-epitope ensemble comprising: SRGWSLIKSVRLGNA, KPRIITLTMNPALDI, AAHKGLMNIALAISA, FPAGGSTGSL, MLLAVTVSL, QSSFYSDW and KMRCGAPRY, with a 97.4% global PPC and a 92.7% East African PPC.

1. INTRODUCTION

Mycobacterium tuberculosis infects populations worldwide. Due in part to troubling rates of new and relapsing tuberculosis (TB), the estimated 2015 death toll from TB was 1.8 million, with 10.5 million new cases recorded (1). Mortality rates are disproportionately high in Africa (2): for example, in 2015, Kenya was reported to have TB affecting over 81,000 people and causing the deaths of at least 16,000 people. TB usually presents as a pulmonary disease transmitted by droplet inhalation resulting in symptoms such as persistent cough, fever, and night sweats (3). Individuals with healthy immune systems can often suppress the disease, typically being asymptomatic but with a latent infection. Problems arise for immunocompromised patients who cannot mount an immune response sufficient for the suppression of symptoms.

The *M. tuberculosis* genome comprises over 4 million base pairs and approximately 4,000 genes (4). The immune response mounted against *M. tuberculosis* mainly involves cellular immunity, including CD4+ and CD8+ T cells (5, 6). Once activated, both CD4+ and CD8+ T cells secrete cytokines inducing an immune response. The CD8+ cells also mediate cytotoxicity and lysis of infected cells. Effective T cell responses are essential for *M. tuberculosis* elimination.

~~However, the~~The slow growth rate of *M. tuberculosis* (792-932 minute doubling time), its complex pathogenesis, and its capacity to remain dormant, are major challenges to the development of effective treatments against TB (7). This is compounded by widespread resistance to antibiotics, such as isoniazid and rifampicin, with 480,000 cases of multi-drug resistant TB emerging annually (8). Thus, vaccination against TB remains a priority; especially in the developing ~~countries~~world, ~~such as~~particularly Africa, India and Indonesia, where the disease is widespread (1).

Commented [GD1]: Africa isn't a country

The *Bacillus Calmette-Guérin* (BCG) vaccine - the only currently available TB vaccine offering prophylaxis - is an attenuated form of *Mycobacterium bovis*, used globally since 1923. Estimates suggest over 3 billion people have been vaccinated with BCG (9). BCG has subsequently come under much scrutiny. Each BCG vaccine dose, whilst containing a preparation of attenuated *M. bovis*, has different biological effects, as the amount of viable versus dead organisms in each dose varies (10). Depending on the strain used, immunogenicity, reactogenicity and viability varies by manufacturer. BCG has efficacy ranging from 0 to 80% against adult pulmonary TB (10, 11), providing protection for 10 to 20 years from immunisation. As a partially-effective vaccine, BCG only protects paediatric patients from severe TB. BCG, as a live-attenuated vaccine, is seen as having a low safety profile due to the risks associated with its use in immune-compromised people and the possibility of the

bacterium reverting to its virulent form (12). BCG's varying efficacy is a particular problem in developing countries, such as East Africa. The need for effective, safe TB vaccines is clear.

Sixteen TB vaccine candidates are currently in phase I, II, and III (13). Almost half of these candidate vaccines contain a live attenuated form of *M. tuberculosis*. Such an approach comes with many safety concerns. Indeed, BCG is contraindicated in patients with human immunodeficiency virus (HIV). Immunoe compromised people are at increased risk of disseminated BCG disease (1). Disseminated BCG disease is a rare, life-threatening complication presenting with symptoms such as persistent fever, infection, weight loss and night sweats (14).

Commented [GD2]: Immune compromised vs immunocompromised. Both are fine but consistency of use needed.

Viral vector-based vaccines include recombinant Vaccinia strain Ankara virus (MVA), synthesised to express antigen 85A of *M. tuberculosis* and an adenovirus expressing the mycobacterial antigens 85A, 85B and TB10.4 known as the Crucell-Ad35/AERAS-402 vaccine (15). However, previous exposure to the vector can reduce vaccine efficacy. Prophylactic live vaccines are also being developed, such as recombinant BCG VPM1002 and MTBVAC (5, 6). Live vaccines can revert to a pathogenic form. This is particularly dangerous for immunocompromised individuals. Inoculating patients with subunit-based vaccines eliminates risks of reversion to virulence. Subunit vaccines include the H1 vaccine which combines 85B and ESAT-6 antigens, H4 which combines antigens 85B and TB10.4 and M72 combining antigens 39A and 32A. However, subunit vaccines lack intrinsic immunogenicity, and are seldom able to induce long-term immunity against diseases, necessitating multiple vaccinations and the inclusion of adjuvants.

Commented [GD3]: See previous comment

The Phase I trial of AERAS-422, a recombinant BCG vaccine, studies, were discontinued when two participants developed the Varicella-Zoster virus (16). M72/AS01$_E$ is a candidate subunit vaccine initially deemed clinical safe in both healthy and TB-infected adults. However, during Stage II trials, many volunteers suffered local reactions at injection sites ending the study prematurely (17). Another subunit vaccine, MVA85A, initially looked promising. Proposed as an adjunct to conventional BCG vaccination, it showed effective protection in animal models. Unfortunately, MVA85A did not show similar efficacy in healthy infants and adults (18).

Given the success of peptide subunit vaccine candidates such as H4/IC31, peptide vaccines should be considered strong potential TB vaccines. H4/IC31 had clinically safe profiles in Phase I trials inducing a positive immune response in healthy adults and infants already vaccinated with BCG (19). Peptide vaccines can be freeze-dried maintaining stability without a cold-chain (20). Apart from low manufacturing costs (21), peptide vaccines typically also have higher safety profiles, due to the use of epitopes without reactogenic responses (20).

Commented [GD4]: This is the first time we mention IC31.

~~Here~~In this work ~~we seek to extend an evolving approach in computational vaccinology by exploring how~~ we sought to design an anti-TB epitope ensemble vaccine applying an ~~evolving~~emerging approach in computational vaccinology. An~~An~~ ideal vaccine would concentrate on highly conserved immunogenic epitopes with a wide population coverage. We have recently begun to exemplify the approach by identifying epitope ensemble vaccine against Hepatitis C (22), influenza (23), and malaria (24). ~~Here~~Likewise, here we ~~seek to select~~selected *M. tuberculosis* epitopes of proven immunogenicity that can be combined to form an effective, widely-applicable epitope ensemble vaccine, defining both a global vaccine and one focussing on East Africa.

Commented [GD5]: Evolving has too many other connotations – like we are using genetic algorithms etc

2. METHODS

2.1. Collection of *Mycobacterium tuberculosis*-Specific Epitopes

The Immune Epitope Database and Analysis Resource (IEDB; URL: http://www.iedb.org/) (25, 26) was used to collect *M. tuberculosis*-specific epitopes. Epitope search terms: any disease, antigen ID: Mycobacterium tuberculosis 1773, human host, and positive T cell assays. MHC Class I and II data were collected separately. The number of MHC I and MHC II epitopes were narrowed by selecting antigens with four or more epitopes (CD8+, Class I) or fifteen or more epitopes (CD4+, MHC Class II).

2.2 Protein Sequences and Multiple Sequence Alignments

Protein sequences of selected antigens were retrieved from the NCBI (National Centre for Biotechnology Information; http://www.ncbi.nlm.nih.gov/). For replaced or removed sequences, the most recently updated sequence was used. The retrieved antigen sequences were run against the NCBI Reference Proteins (RefSeq) database using ~~remote~~automated BLASTP (27), with maximum hit sequences limited to 100 and restricting the search to *Mycobacterium tuberculosis*. Multiple Sequence Alignments (MSAs) were generated using COBALT (28).

Formatted: Font: Italic

2.3. Analysis of Epitope Sequence Variability

Conserved epitopes were identified by determining variable versus conserved sequence positions by analysing each separate MSA using the Protein Variability Server, PVS hereafter (29) (URL: http://imed.med.ucm.es/PVS/). Sequence variability was masked and only fragments with a length greater or equal to 9 were selected. The Shannon entropy threshold was set to 0.5.

2.4. Prediction of HLA binding profile for Conserved Epitopes

IEDB was used to calculate the binding profiles of MHC Class I (http://tools.iedb.org/mhci/) and Class II (http://tools.iedb.org/mhcii/) highly conserved epitopes. HLA I reference set was used for MHC I epitopes (30) and an HLA II reference set was used for MHC II epitopes (31). For MHC Class I, epitopes were chosen with a percentile rank less than or equal to one. For MHC Class II, epitopes with a percentile rank greater than or equal to five were omitted.

2.5. Calculation of Projected Population Coverage (PPC) and Optimal Epitope Identification

The PPC of highly conserved epitopes were calculated using IEDB (http://tools.iedb.org/tools/population/iedb_input). MHC I and MHC II epitopes were then ranked by PPC. Epitopes were combined within each class to calculate PPC's for the World and East Africa.

MHC I and II epitopes were also combined and the PPC calculated. Vaccine combinations were ranked by PPC's in descending order with criteria of 90% or higher for the world and 80% or higher for East Africa.

3. RESULTS

3.1 MHC Class I Epitope selection and combination

Using IEDB, a total of 400 epitopes were obtained from 122 different TB antigens. Following Damfo, Reche (24), we focussed on epitope-rich antigens containing four or more epitopes: 18 different antigens were identified containing a total of 259 epitopes (Figure 1A). Of these, 151 were conserved, as determined by PVS (29). The binding profile and percentage population coverage of each epitope was calculated using IEDB. Epitopes were selected on the allele diversity of their binding profiles and a PPC of at least 20% for the World or East Africa (Table 1).

Commented [GD6]: I realise EndNote has done this – but is this really the journal's preferred format?

By choosing the epitope with the highest global PPC (MLLAVTVSL) and then selecting epitopes that bound different HLA I alleles, PPC values for epitope combinations were calculated. Combining MHC I epitopes - QSSFYSDW, ALAALGLWLSV, MLLAVTVSL, AEQEQCLDEL, FPAGGSTGSL, HISSGVFLLK, KMRCGAPRY, WYYQSGLSI, NTPAFEWYY - gave a maximum PPC of 97.8% for the World and 86.1% for East Africa. A comparable combination - QSSFYSDW, MLLAVTVSL, AEQEQCLDEL, FPAGGSTGSL, HISSGVFLLK, KMRCGAPRY, WYYQSGLSI - gave a PPC of 97.1% for the World and 85.6% for East Africa.

3.2 MHC Class II Epitope selection and combination

Using IEDB, a total of 1007 epitopes were obtained from 325 different TB antigens. By focussing on antigens with fifteen or more epitopes, 15 different antigens were identified. (Figure 1B). Of the resulting 381 epitopes, 239 were conserved, as determined using PVS. The binding profile and percentage population coverage of each epitope was calculated using IEDB. Epitopes were selected on the allele diversity of their binding profiles and a PPC of at least 20% for the World or East Africa (Table 2).

By choosing the epitope with the highest global PPC (KPRIITLTMNPALDI) and then selecting epitopes that bound different HLA I alleles, PPC values for epitope combinations were calculated. The MHC II epitope combination comprising SRGWSLIKSVRLGNA, KPRIITLTMNPALDI and AAHKGLMNIALAISA gave a PPC of 81.8% for the World and 68.3% for East Africa.

3.3 MHC I and II Combined

To increase the overall PPC values, and achieve PPC over 75% in East Africa and 85% globally, MHC I and II epitopes were combined (Table 3). All combinations contained the final set of MHC II epitopes. Given the constraint of providing the greatest projected population coverage with the minimum number of epitopes, we did not combine all MHC I and MHC II epitopes. Effective PPC's were considered to have greater than 90% coverage for the World and greater than 80% coverage for East Africa (Table 3).

3.4 Cross-reactivity with the Human Proteome

As a surrogate for potentially dangerous auto-immune reaction, ~~ADRs, etc~~ mediated by epitope cross-reactivity with the human proteome, we compared the sequences of the final epitope selection comprising the highest scoring vaccine with BLASTP, which automatically adjusts for short sequence lengths. The results - SRGWSLIKSVRLGNA [best overlap: 60%; % identity: 64%; best match: AAK97224.1], KPRIITLTMNPALDI [60%; 78%; BAF94735.1], AAHKGLMNIALAISA [46%; 100%; EAW74100.1], FPAGGSTGSL [80%; 100%; XP_016857235.1], MLLAVTVSL [100%; 89%; EAW77954.1], QSSFYSDW [75%; 100%; XP_016863485.1], KMRCGAPRY [66%; 100%; CAC51389.1] – indicate, at best, poor partial matches to a variety of human proteins.

Commented [GD7]: This is the first mention of ADRs

4. DISCUSSION

We have sought to define a putative epitope ensemble vaccine selecting epitopes of known immunogenicity to optimise predicted population coverage. Although the BCG vaccine has been used for many years, it has a varying efficacy especially in adults presenting with pulmonary TB. A viable TB vaccine should have high specificity with universal coverage to reduce the rate of disease-associated illness and death. Our aim was to identify combinations of conserved epitopes able to stimulate appropriate, desirable immune responses against *M. tuberculosis*. We generated a number of essentially equivalent combinations of class I and class II epitopes to yield a PPC above 90% (universal or world vaccine) and above 80% for an East Africa-specific vaccine. The best MHC II epitope combination gave a maximum PPC lower than expected, although typically MHC II epitopes have a greater range of HLA binding profiles and should thus have higher PPC values.

Vaccines should induce immune responses of a quality and strength sufficient to induce effective protection (32) without immune-mediated adverse reactions, such as anaphylaxis. Peptide vaccines can do this by using an optimal number of epitopes: the fewer the epitopes used, the less likely negative immunogenicity. Moreover, potential adverse reactions may also be mitigated by focussing the immune response to a restricted set of antigens, since side reactions can likewise be minimised.

Of the twelve epitopes that made it into the final combinations of potential vaccines, six (KPRIITLTMNPALDI, SRGWSLIKSVRLGNA, AEQEQCLDEL, FPAGGSTGSL, HISSGVFLLK and KMRCGAPRY) come from phosphofructokinase PfkB, an essential gene in TB, a vital enzyme in carbohydrate metabolism. It has low similarity to host sequences, reducing the risk of auto-immunity (33). Two of the final epitopes – ALAALGLWLSV and MLLAVTVSL - come from the gene Rv1733c, often described as a latent antigen. It produces a high T cell response in latently infected patients and is highly expressed by dormant versions of the bacterium.

Epitopes WYYQSGLSI and NTPAFEWYY arise from the *M. tuberculosis* gene fbpB, involved in the maintenance of the Mycobacterial cell wall (34). They originate from the secreted antigen 85-B, present during early infection. Also known as Rv1886c, antigen 85-B has long been viewed as a potential subunit vaccine candidate (35). Expression of Ag85-B is greater than that of its counterparts Ag85-A and Ag85-C, being a key secreted proteins of *M. tuberculosis* (36). QSSFYSDW derives from antigen 85-A, which is currently undergoing clinical trials as a candidate subunit vaccine (37). The MHC Class II epitope AAHKGLMNIALAISA, derives from a phosphate-binding protein which can stimulate CD4+ T cells, leading to production of IFN-γ, IL-22 and IL-17 (38).

In terms of HLA I diversity versus pathogen-prevalence, African populations have greater HLA diversity (39). Sub-Saharan Africa ~~may be~~is the most genetically-diverse population in ~~Africa~~the world. (40) studied the commonest HLA I alleles in the East Africa: HLA-A*01:01, HLA-A*02:01, HLA-A*03:01, HLA-A*23:01, HLA-A*30:02, HLA-B*07:02, HLA-B*08:01, HLA-B*35:01, HLA-B*44:03, HLA-B*51:01, HLA-B*53:01, HLA-B*57:01 and HLA-B*58:01 cover 80 to 90% of the East African population. All these alleles are found in our final epitope combinations. The FPAGGSTGSL MHC I epitope, with an East African PPC value of 33.6%, binds to all the common alleles in East Africa. The same applies for the QSSFYSDW epitope which has the highest East African population coverage (34.1%).

Dunand, Ng (41) highlights the genetic differences between black East Africans and other African populations, suggesting the need to create a TB vaccine specifically targeting East African. Of the HLA II genes, the HLA-DRB1 alleles exhibits the highest polymorphism (41). Chen et al. (2015) have shown this allele to have a likely protective role against *M. tuberculosis*. Dunand et al. (1997) showed HLA-DRB1*11:01 occurs frequently in the Kenyan population. The three MHC II epitopes used in the final vaccine combinations (SRGWSLIKSVRLGNA, KPRIITLTMNPALDI and AAHKGLMNIALAISA), all bind this allele leading to high East African PPC values.

The final vaccine combinations typically had higher predicted population coverages for the World than East Africa. Combinations were chosen to optimise PPC (World > 90%, East Africa > 80%) and to cover range of allele binding profiles while using an optimal number of epitopes. Making use of a small number of epitopes increases the likely occurrence of only positive immunogenicity. By choosing epitopes with wide-ranging HLA I and II binding profiles, higher intrinsic and cumulative population coverage are guaranteed. Of the five vaccines, three arose by combining CD8+ and CD4+ epitopes.

Antigens with a greater number of epitopes were more likely to produce highly conserved and effective epitopes. For the MHC I Class, the highest antigens per epitope were NP_216545.1, NP_216249.1 and YP_001283215.1 (Figure 1). These antigens contained eight epitopes which were highly conserved and offered effective protection against Tuberculosis. Similarly, NP_216545.1 had the second highest epitope count. It contained two highly-conserved epitopes included in the final epitope combinations.

CD4+ epitopes are typically longer and more prevalent, and potentially bind many more alleles than CD8+ epitopes. However, CD4+ combinations produced PPC values lower than expected. The low PPC values of 81.8% (World) and 68.3% (East Africa) might be raised if more antigens were used. Antigens with 15 or more CD4+ epitopes were used for logistic reasons and to

focus the resulting immune response; setting a lower value – say 10 epitopes – might offer a catholic selection without broadening the response unduly.

Our vaccine does not contain known B-cell epitopes or mimics thereof. Incorporating B cell antigens into TB vaccines has been tried in the formulation of the CTA1-DD/ISCOM adjuvant vector. Fusing this vector with ESAT-6/Ag85B boosted BCG induced immunity (42). However, a major issue in B cell epitope identification is the total lack of accuracy and thus low success rate in their prediction (43, 44). Given the dominance of cellular immunity in responses to *M. tuberculosis*, B cell epitopes were deprecated in favour of their T cell equivalent.

~~As a surrogate for potentially dangerous immune reaction,~~ Again, we compared the sequences comprising the highest scoring vaccine to the human virtual proteome. Even the best match - MLLAVTVSL [overlap: 100%; %identity: 89%; best match: EAW77954.1] – is a partial match. It is though that multiple complete matches between an epitope and distinct human proteins are necessary for such correspondence to become an issue (45,46,47). As the number of complete matches rises, the chance that the native epitope will be presented as part of the human "self" peptidome likewise rises, the necessary condition for epitope-mediated immune reaction. As more peptidomes are sequenced, the more reliable such predictions will become.

5. CONCLUSION

We used an immunoinformatic protocol to define a vaccine targeting *Mycobacterium tuberculosis*, producing a combined MHC I and MHC II epitope ensemble vaccine that offered 97.4% global protection and 92.67% coverage for the East African population. The final seven-epitope combination comprised: SRGWSLIKSVRLGNA, KPRIITLTMNPALDI, AAHKGLMNIALAISA, FPAGGSTGSL, MLLAVTVSL, QSSFYSDW and KMRCGAPRY; indicate that our putative TB vaccines need both MHC I and II epitopes to offer a broad population coverage. *In vitro* and *in vivo* testing of our ensemble vaccine for protective immune responses and efficacy against *Mycobacterium tuberculosis*. This and other work (22-24) indicates our evolving immunoinformatics design strategy, based on the rigorous selection of pre-validated epitopes, is suitable for developing epitope ensemble vaccines across the spectrum of key pathogens for human health.

Formatted: Font: 11 pt

Formatted: Font: 11 pt

Formatted: Font: 11 pt

Conflicts of Interest

There are no conflicts to declare

Acknowledgements

Funding: This research was supported by Aston University. PAR is supported by grants BIO2014:54164-R and by Inmunotek S.L.

Formatted: Font: 11 pt

REFERENCES

1. WHO. Tuberculosis fact sheet, [online]. http://www.who.int/mediacentre/factsheets/fs104/en/ (accessed: 15th March 17)2017.
2. Migliori GB, Dheda K, Centis R, Mwaba P, Bates M, O'Grady J, et al. Review of multidrug-resistant and extensively drug-resistant TB: global perspectives with a focus on sub-Saharan Africa. Trop Med Int Health. 2010;15(9):1052-66.
3. Heemskerk D, Caws M, Marais B, Farrar J. Tuberculosis in Adults and Children. Wellcome Trust–Funded Monographs and Book Chapters. London2015.
4. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, et al. Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. Nature. 1998;393(6685):537-44.
5. Mendez-Samperio P. Global Efforts in the Development of Vaccines for Tuberculosis: Requirements for Improved Vaccines Against Mycobacterium tuberculosis. Scandinavian Journal of Immunology. 2016;84(4):204-10.
6. Mendez-Samperio P. Research progress in the field of immunotherapeutic vaccination in human TB and the way ahead. Immunotherapy. 2016;8(9):987-9.
7. Zenteno-Cuevas R. Successes and failures in human tuberculosis vaccine development. Expert Opin Biol Ther. 2017:1-11.
8. Raviglione M, Sulis G. Tuberculosis 2015: Burden, Challenges and Strategy for Control and Elimination. Infect Dis Rep. 2016;8(2):6570.
9. Franco-Paredes C, Rouphael N, del Rio C, Santos-Preciado JI. Vaccination strategies to prevent tuberculosis in the new millennium: from BCG to new vaccine candidates. Int J Infect Dis. 2006;10(2):93-102.
10. Bali P, Tousif S, Das G, Van Kaer L. Strategies to improve BCG vaccine efficacy. Immunotherapy. 2015;7(9):945-8.
11. Romano M, Huygen K. An update on vaccines for tuberculosis - there is more to it than just waning of BCG efficacy with time. Expert Opin Biol Ther. 2012;12(12):1601-10.
12. Detmer A, Glenting J. Live bacterial vaccines--a review and identification of potential hazards. Microb Cell Fact. 2006;5:23.
13. Evans TG, Schrager L, Thole J. Status of vaccine research and development of vaccines for tuberculosis. Vaccine. 2016;34(26):2911-4.
14. Eccles SR, Mehta R. Disseminated BCG disease: A case report. Respiratory Medicine CME. 2011;4(3):112-3.
15. Wilkie ME, McShane H. TB vaccine development: where are we and why is it so difficult? Thorax. 2015;70(3):299-301.
16. Hoft DF, Blazevic A, Selimovic A, Turan A, Tennant J, Abate G, et al. Safety and Immunogenicity of the Recombinant BCG Vaccine AERAS-422 in Healthy BCG-naive Adults: A Randomized, Active-controlled, First-in-human Phase 1 Trial. EBioMedicine. 2016;7:278-86.
17. Gillard P, Yang PC, Danilovits M, Su WJ, Cheng SL, Pehme L, et al. Safety and immunogenicity of the M72/AS01E candidate tuberculosis vaccine in adults with tuberculosis: A phase II randomised study. Tuberculosis (Edinb). 2016;100:118-27.
18. Ndiaye BP, Thienemann F, Ota M, Landry BS, Camara M, Dieye S, et al. Safety, immunogenicity, and efficacy of the candidate tuberculosis vaccine MVA85A in healthy adults infected with HIV-1: a randomised, placebo-controlled, phase 2 trial. Lancet Respir Med. 2015;3(3):190-200.
19. Kagina BM, Tameris MD, Geldenhuys H, Hatherill M, Abel B, Hussey GD, et al. The novel tuberculosis vaccine, AERAS-402, is safe in healthy infants previously vaccinated with BCG, and induces dose-dependent CD4 and CD8T cell responses. Vaccine. 2014;32(45):5908-17.
20. Li W, Joshi MD, Singhania S, Ramsey KH, Murthy AK. Peptide Vaccine: Progress and Challenges. Vaccines (Basel). 2014;2(3):515-36.
21. Slingluff CL, Jr. The present and future of peptide vaccines for cancer: single or multiple, long or short, alone or in combination? Cancer J. 2011;17(5):343-50.

22. Molero-Abraham M, Lafuente EM, Flower DR, Reche PA. Selection of conserved epitopes from hepatitis C virus for pan-populational stimulation of T-cell responses. Clin Dev Immunol. 2013;2013:601943.
23. Sheikh QM, Gatherer D, Reche PA, Flower DR. Towards the knowledge-based design of universal influenza epitope ensemble vaccines. Bioinformatics. 2016;32(21):3233-9.
24. Damfo S, Reche P, Gatherer D, Flower D. In silico design of knowledge-based Plasmodium falciparum epitope ensemble vaccines. Journal of Molecular Graphics and Modelling. 2017:in press.
25. Peters B, Sidney J, Bourne P, Bui HH, Buus S, Doh G, et al. The immune epitope database and analysis resource: from vision to blueprint. PLoS Biol. 2005;3(3):e91.
26. Peters B, Sidney J, Bourne P, Bui HH, Buus S, Doh G, et al. The design and implementation of the immune epitope database and analysis resource. Immunogenetics. 2005;57(5):326-36.
27. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997;25(17):3389-402.
28. Papadopoulos JS, Agarwala R. COBALT: constraint-based alignment tool for multiple protein sequences. Bioinformatics. 2007;23(9):1073-9.
29. Garcia-Boronat M, Diez-Rivero CM, Reinherz EL, Reche PA. PVS: a web server for protein sequence variability analysis tuned to facilitate conserved epitope discovery. Nucleic Acids Res. 2008;36(Web Server issue):W35-41.
30. Weiskopf D, Angelo MA, de Azeredo EL, Sidney J, Greenbaum JA, Fernando AN, et al. Comprehensive analysis of dengue virus-specific responses supports an HLA-linked protective role for CD8+ T cells. Proc Natl Acad Sci U S A. 2013;110(22):E2046-53.
31. Greenbaum J, Sidney J, Chung J, Brander C, Peters B, Sette A. Functional classification of class II human leukocyte antigen (HLA) molecules reveals seven different supertypes and a surprising degree of repertoire sharing across supertypes. Immunogenetics. 2011;63(6):325-35.
32. Kaufmann SH, McElrath MJ, Lewis DJ, Del Giudice G. Challenges and responses in human vaccine development. Curr Opin Immunol. 2014;28:18-26.
33. Amir A, Rana K, Arya A, Kapoor N, Kumar H, Siddiqui MA. Mycobacterium tuberculosis H37Rv: In Silico Drug Targets Identification by Metabolic Pathways Analysis. Int J Evol Biol. 2014;2014:284170.
34. Davila J, Zhang L, Marrs CF, Durmaz R, Yang Z. Assessment of the genetic diversity of Mycobacterium tuberculosis esxA, esxH, and fbpB genes among clinical isolates and its implication for the future immunization by new tuberculosis subunit vaccines Ag85B-ESAT-6 and Ag85B-TB10.4. J Biomed Biotechnol. 2010;2010:208371.
35. Zvi A, Ariel N, Fulkerson J, Sadoff JC, Shafferman A. Whole genome identification of Mycobacterium tuberculosis vaccine candidates by comprehensive data mining and bioinformatic analyses. BMC Med Genomics. 2008;1:18.
36. Aghababa H, Mohabati Mobarez A, Khoramabadi N, Behmanesh M, Mahdavi M, Tebianian M, et al. A comparative approach to strategies for cloning, expression, and purification of Mycobacterium tuberculosis mycolyl transferase 85B and evaluation of immune responses in BALB/c mice. Mol Biotechnol. 2014;56(6):487-97.
37. Smaill F, Xing Z. Human type 5 adenovirus-based tuberculosis vaccine: is the respiratory route of delivery the future? Expert Rev Vaccines. 2014;13(8):927-30.
38. Palma C, Schiavoni G, Abalsamo L, Mattei F, Piccaro G, Sanchez M, et al. Mycobacterium tuberculosis PstS1 amplifies IFN-gamma and induces IL-17/IL-22 responses by unrelated memory CD4+ T cells via dendritic cell activation. Eur J Immunol. 2013;43(9):2386-97.
39. Prugnolle F, Manica A, Charpentier M, Guegan JF, Guernier V, Balloux F. Pathogen-driven selection and worldwide HLA class I diversity. Curr Biol. 2005;15(11):1022-7.
40. Koehler RN, Walsh AM, Sanders-Buell EE, Eller LA, Eller M, Currier JR, et al. High-throughput high-resolution class I HLA genotyping in East Africa. PLoS One. 2010;5(5):e10751.
41. Dunand VA, Ng CM, Wade JA, Bwayo J, Plummer FA, MacDonald KS. HLA-DR 52- and 51-associated DRB1 alleles in Kenya, east Africa. Tissue Antigens. 1997;49(4):397-402.
42. Andersen CS, Dietrich J, Agger EM, Lycke NY, Lovgren K, Andersen P. The combined CTA1-DD/ISCOMs vector is an effective intranasal adjuvant for boosting prior Mycobacterium bovis BCG immunity to Mycobacterium tuberculosis. Infect Immun. 2007;75(1):408-16.
43. Blythe MJ, Flower DR. Benchmarking B cell epitope prediction: underperformance of existing methods. Protein Sci. 2005;14(1):246-8.
44. Ponomarenko JV, Bourne PE. Antibody-protein interactions: benchmark datasets and prediction tools evaluation. BMC Struct Biol. 2007;7:64.

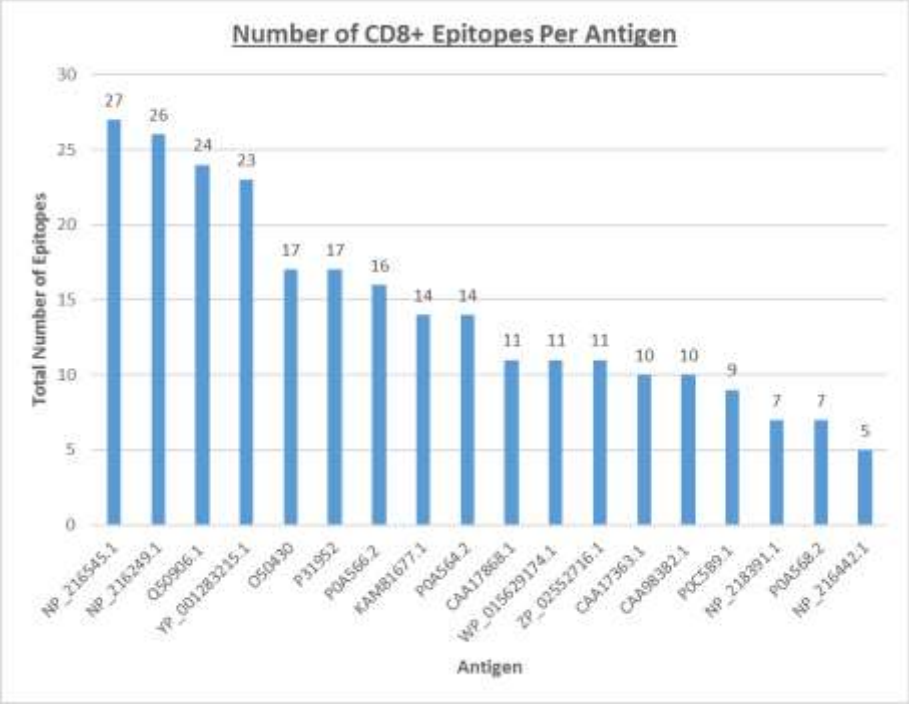
45. Ramakrishnan K, Flower DR. Discriminating antigen and non-antigen using proteome dissimilarity: bacterial antigens. Bioinformation. 2010; 4(10):445-447.
46. Ramakrishnan K, Flower DR. Discriminating antigen and non-antigen using proteome dissimilarity III: tumour and parasite antigens. Bioinformation. 2010; 5(1):39-42.
47: Ramakrishnan K, Flower DR. Discriminating antigen and non-antigen using proteome dissimilarity II: viral and fungal antigens. Bioinformation. 2010; 24;5(1):35-38.

FIGURE 1: Prevalence of Epitopes within TB Antigens

Formatted: Font: 11 pt

Prevalence of Epitopes within TB Antigens. A: Class I. B. Class II.

A. The height of each bar represents the number of epitopes present in the named antigen. For completeness and accuracy, the number of epitopes also appears above the bar. 18 antigens contained 5 or more epitopes.



B. The height of each bar represents the number of epitopes present in the named antigen. For completeness and accuracy, the number of epitopes also appears above the bar. 14 antigens contained 15 or more epitopes.

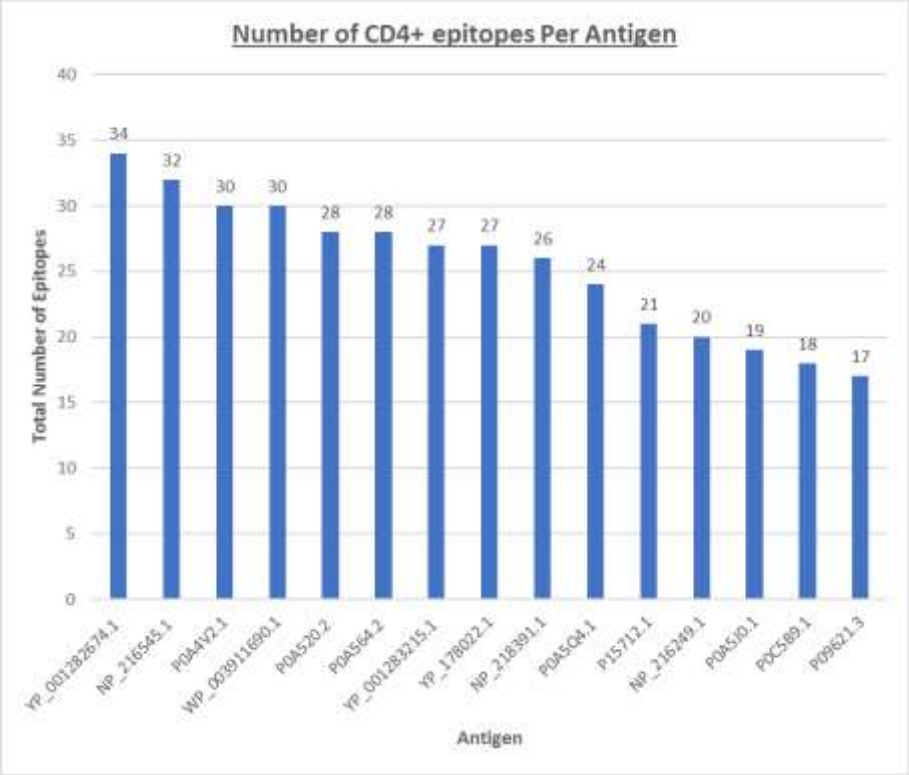


TABLE 1: Characteristics of Conserved CD8+ Epitopes

This Table lists the sequences of the conserved CD8+ epitopes identified by the PVS server, as ranked in terms of the overall Projected Population Coverage or PPC, given as a percentage, for both the world and for East Africa. The PPC is calculated from the prevalence of alleles predicted as binding to the given epitope, as listed in the table. The Tuberculosis antigen of origin is also given.

EPITOPE	ANTIGEN ID	MHC class I BINDING PROFILE	PROJECTED POPULATION COVERAGE (%)	
			WORLD	EAST AFRICA
MLLAVTVSL	NP_216249.1	HLA-A*32:01, HLA-A*02:01, HLA-B*08:01, HLA-A*02:06	50.03	28.34
ALAALGLWLSV	NP_216249.1	HLA-A*02:01, HLA-A*02:03, HLA-A*02:06, HLA-A*32:01	44.86	23.04
FVLPGPSLTV	NP_216545.1	HLA-A*02:06, HLA-A*02:01, HLA-A*02:03	41.35	20.47
AAMASASLV	Q50906.1	HLA-A*02:01	39.08	20.24
FPAGGSTGSL	NP_216545.1	HLA-B*07:02, HLA-B*53:01, HLA-B*35:01, HLA-B*51:01, HLA-B*08:01	38.21	33.60
HISSGVFLLK	NP_216545.1	HLA-A*11:01, HLA-A*03:01, HLA-A*68:01	35.75	11.72
KAGCQTYK	YP_001283215.1	HLA-A*30:01, HLA-A*03:01, HLA-A*11:01	34.14	20.40
AVVRFQEAANK	POA566.2	HLA-A*11:01, HLA-A*03:01	30.92	9.20
KMRCGAPRY	NP_216545.1	HLA-A*30:02, HLA-B*15:01, HLA-A*30:01, HLA-A*03:01	29.06	32.22
YQGVQQKW	P0A564.2	HLA-A*24:02, HLA-B*58:01, HLA-B*53:01, HLA-B*57:01	29.00	27.77
WYYQSGLSI	YP_001283215.1	HLA-A*23:01, HLA-A*24:02	26.18	17.08
GVAADYYQR	NP_216545.1	HLA-A*31:01, HLA-	25.64	6.47

Formatted: Font: 11 pt, English (United Kingdom)
Formatted: Font: 11 pt, English (United Kingdom)
Formatted: Font: 11 pt
Formatted: Font: 11 pt, English (United Kingdom)
Formatted: Font: 11 pt, English (United Kingdom)
Formatted: Font: 11 pt
Formatted: Font: 11 pt, English (United Kingdom)
Formatted: Font: 11 pt, English (United Kingdom)
Formatted: Font: 11 pt, English (United Kingdom)
Formatted: Font: 11 pt, English (United Kingdom)
Formatted: Font: 11 pt, English (United Kingdom)
Formatted: Font: 11 pt, English (United Kingdom)
Formatted: Font: 11 pt, English (United Kingdom)
Formatted: Font: 11 pt, English (United Kingdom)
Formatted: Font: 11 pt
Formatted: Font: 11 pt, English (United Kingdom)
Formatted: Font: 11 pt, English (United Kingdom)
Formatted: Font: 11 pt, English (United Kingdom)

		A*68:01, HLA-A*11:01		
QSSFYSDW	CAA17868.1	HLA-B*58:01, HLA-B*57:01, HLA-A*01:01, HLA-A*30:02 HLA-A*01:01, HLA-B*15:01, HLA-B*15:01	25.38	34.08
FSRPGLPVEY	CAA17868.1	HLA-A*01:01, HLA-B*15:01, HLA-B*15:01	24.32	12.36
MPVGGQSSF	CAA17868.1	HLA-B*35:01, HLA-B*53:01, HLA-B*07:02, HLA-B*35:01, HLA-B*53:01	22.88	23.90
MAEMKTDAATL	NP_218391.1	HLA-B*40:01, HLA-B*44:02, HLA-B*15:01	22.88	3.37
NTPAFEWYY	YP_001283215.1	HLA-A*01:01, HLA-A*26:01	22.62	13.04
VSLLTIPF	NP_216249.1	HLA-A*32:01, HLA-A*32:01, HLA-B*57:01, HLA-B*57:01, HLA-A*30:01, HLA-A*30:01, HLA-B*58:01, HLA-B*58:01, HLA-B*15:01, HLA-B*15:01	22.49	26.44
AEQEQCLDEL	NP_216545.1	HLA-B*40:01, HLA-B*44:02, HLA-B*44:03	20.88	6.94
RGAAASAAF	NP_216545.1	HLA-B*15:01, HLA-B*07:02	20.65	7.37
LPVEYLQV	CAA17868.1	HLA-B*07:02, HLA-B*07:02, HLA-B*35:01, HLA-B*35:01	20.62	10.51
AVYLLDGLR	YP_001283215.1	HLA-A*31:01, HLA-A*11:01	20.45	3.87
FAAAAGTAV	NP_216249.1	HLA-A*02:06, HLA-A*68:02, HLA-B*35:01, HLA-A*02:03, HLA-B*51:01	20.07	22.73
QSSFYSDWY	P31952	HLA-A*01:01, HLA-A*30:02	19.55	24.39
SSFYSDWY	CAA17868.1	HLA-B*57:01, HLA-	17.44	26.02

Formatted: Font: 11 pt, English (United Kingdom)

Formatted: Font: 11 pt, English (United Kingdom)

Formatted: Font: 11 pt

Formatted: Font: 11 pt, English (United Kingdom)

Formatted: Font: 11 pt, English (United Kingdom)

Formatted: Font: 11 pt

Formatted: Font: 11 pt, English (United Kingdom)

Formatted: Font: 11 pt, English (United Kingdom)

Formatted: Font: 11 pt

Formatted: Font: 11 pt, English (United Kingdom)

Formatted: Font: 11 pt, English (United Kingdom)

Formatted: Font: 11 pt, English (United Kingdom)

Formatted: Font: 11 pt

Formatted: Font: 11 pt, English (United Kingdom)

Formatted: Font: 11 pt, English (United Kingdom)

Formatted: Font: 11 pt

Formatted: Font: 11 pt, English (United Kingdom)

Formatted: Font: 11 pt, English (United Kingdom)

Formatted: Font: 11 pt, English (United Kingdom)

Formatted: Font: 11 pt

Formatted: Font: 11 pt, English (United Kingdom)

Formatted: Font: 11 pt, English (United Kingdom)

Formatted: Font: 11 pt, English (United Kingdom)

Formatted: Font: 11 pt

Formatted: Font: 11 pt, English (United Kingdom)

Formatted: Font: 11 pt, English (United Kingdom)

Formatted: Font: 11 pt, English (United Kingdom)

Formatted: Font: 11 pt

Formatted: Font: 11 pt, English (United Kingdom)

		B*58:01, HLA-B*15:01, HLA-A*30:02		
VGQDQYVW	NP_216545.1	HLA-B*57:01, HLA-B*57:01, HLA-B*58:01, HLA-B*58:01, HLA-B*53:01, HLA-B*53:01	9.70	26.01

Formatted: Font: 11 pt, English (United Kingdom)

TABLE 2: Characteristics of Conserved CD4+ Epitopes

This Table lists the sequences of the conserved CD4+ epitopes identified by the PVS server, as ranked in terms of the overall Projected Population Coverage or PPC, given as a percentage, for both the world and for East Africa. The PPC is calculated from the prevalence of alleles predicted as binding to the given epitope, as listed in the table. The Tuberculosis antigen of origin is also given.

EPITOPE	ANTIGEN ID	MHC class II BINDING PROFILE	PROJECTED POPULATION COVERAGE (%)	
			World	East Africa
KPRIITLTMNPALDI	NP_216545.1	HLA-DRB4*01:01, HLA-DRB1*13:02, HLA-DRB1*04:05, HLA-DRB1*09:01, HLA-DRB1*11:01, HLA-DRB1*04:01, HLA-DRB3*02:02, HLA-DRB1*15:01, HLA-DRB1*08:02, HLA-DRB1*07:01, HLA-DRB1*03:01	74.13	59.68
LEAVVMLLAVTVSLL	NP_216249.1	HLA-DRB1*07:01, HLA-DRB1*03:01, HLA-DRB1*11:01, HLA-DRB1*15:01, HLA-DRB1*04:01, HLA-DRB1*04:05, HLA-DPA1*03:01/DPB1*04:02, HLA-DQA1*01:02/DQB1*06:02, HLA-DRB1*08:02, HLA-DRB5*01:01, HLA-DRB1*01:01	73.40	51.56
SRGWSLIKSVRLGNA	NP_216545.1	HLA-DRB1*07:01, HLA-DRB1*08:02, HLA-DRB1*01:01, HLA-DRB3*02:02, HLA-DRB5*01:01, HLA-DRB1*04:01, HLA-DRB1*04:05, HLA-DRB1*15:01, HLA-DRB1*11:01, HLA-DRB1*03:01	73.40	51.56
AELMILIATNLLGQN	YP_001282674.1	HLA-DRB1*04:01, HLA-DRB1*01:01, HLA-DRB1*15:01, HLA-DRB1*12:01, HLA-DRB1*04:05, HLA-DRB5*01:01, HLA-DPA1*03:01/DPB1*04:02, HLA-DRB1*03:01, HLA-	67.91	61.06

Formatted: Font: 11 pt

Formatted: Font: 11 pt

Formatted: Font: 11 pt

Formatted: Font: 11 pt

Formatted: Font: 11 pt

		DRB1*11:01, HLA-DRB1*13:02		
AAVVMLLAVTVSLLTI	NP_216249.1	HLA-DRB1*07:01, HLA-DRB1*03:01, HLA-DRB1*11:01, HLA-DRB1*04:01, HLA-DRB1*04:05, HLA-DPA1*03:01/DPB1*04:02, HLA-DPA1*01/DPB1*04:01, HLA-DRB1*08:02, HLA-DRB1*15:01, HLA-DRB5*01:01	66.92	48.73
ATNFFGINTIPIALF	YP_178022.1	HLA-DRB1*04:01, HLA-DRB1*11:01, HLA-DRB1*04:05, HLA-DRB1*08:02, HLA-DRB1*07:01, HLA-DRB1*13:02, HLA-DQA1*01:01/DQB1*05:01, HLA-DRB5*01:01, HLA-DRB1*12:01, HLA-DPA1*02:01/DPB1*14:01, HLA-DPA1*02:01/DPB1*01:01, HLA-DRB3*02:02, HLA-DRB1*03:01	62.15	64.00
AALAIAAMASASLVT	WP_003911690.1	HLA-DRB1*09:01, HLA-DQA1*05:01/DQB1*03:01, HLA-DRB1*01:01, HLA-DRB1*07:01, HLA-DQA1*01:02/DQB1*06:02, HLA-DRB1*15:01, HLA-DRB1*04:05, HLA-DRB1*08:02, HLA-DRB1*04:01	60.19	19.90
TRAILIRVRNASWQH	NP_216249.1	HLA-DRB1*08:02, HLA-DRB1*15:01, HLA-DRB1*11:01, HLA-DRB1*04:05, HLA-DRB1*04:01, HLA-DRB4*01:01, HLA-DRB1*03:01, HLA-DRB1*13:02	59.44	51.56
AGLMVAAASPYVAWM	YP_001282674.1	HLA-DRB3*01:01, HLA-DRB1*08:02, HLA-DRB1*09:01, HLA-DRB1*12:01, HLA-DRB1*15:01, HLA-DRB1*03:01, HLA-DRB3*02:02, HLA-DRB1*01:01	53.29	29.27
RDVLAVVSKASGLVI	P09621.3	HLA-DRB1*08:02, HLA-DRB1*11:01, HLA-DRB1*07:01, HLA-DRB5*01:01, HLA-DRB1*15:01, HLA-DRB1*13:02	49.93	44.94
AAHKGLMNIALAISA	P15712.1	HLA-DRB1*09:01, HLA-DRB1*11:01, HLA-DRB1*08:02, HLA-DQA1*01:02/DQB1*06:02, HLA-DRB1*12:01, HLA-DRB1*03:01, HLA-DRB1*04:01	47.05	48.30

Formatted: Font: 11 pt

Formatted: Font: 11 pt

Formatted: Font: 11 pt

Formatted: Font: 11 pt

Formatted: Font: 11 pt

Formatted: Font: 11 pt

Formatted: Font: 11 pt

PAFEWYYQSGLSIVM	YP_0012832 15.1	HLA-DRB1*07:01, HLA-DPA1*02:01/DPB1*01:01, HLA-DRB1*11:01, HLA-DRB5*01:01, HLA-DPA1*02:01/DPB1*14:01, HLA-DRB1*04:01, HLA-DQA1*01:01/DQB1*05:01, HLA-DRB1*09:01, HLA-DPA1*01:03/DPB1*02:01, HLA-DPA1*01/DPB1*04:01	42.27	36.80
AFAGIEAAASAIQGN	P0A564.2	HLA-DRB1*13:02, HLA-DRB1*07:01, HLA-DPA1*03:01/DPB1*04:02, HLA-DRB1*03:01, HLA-DRB3*01:01, HLA-DPA1*02:01/DPB1*01:01, HLA-DPA1*02:01/DPB1*14:01	39.68	33.09
AFEWLQVPSPSMGRD	POA4V2.1	HLA-DRB1*04:01, HLA-DRB1*09:01, HLA-DRB1*04:05, HLA-DRB1*01:01, HLA-DRB1*11:01	39.02	36.16
AVYLLDGLRAQSGLS	P0C5B9.1	HLA-DRB1*11:01, HLA-DRB3*01:01, HLA-DRB1*08:02, HLA-DRB1*04:01, HLA-DRB1*03:01	38.72	38.22
CQTYKWETFLTSELP	POA4V2.1	HLA-DPA1*01/DPB1*04:01, HLA-DPA1*02:01/DPB1*05:01, HLA-DPA1*02:01/DPB1*01:01, HLA-DPA1*02:01/DPB1*14:01, HLA-DPA1*03:01/DPB1*04:02, HLA-DRB1*07:01, HLA-DQA1*01:02/DQB1*06:02, HLA-DRB1*08:02, HLA-DRB1*03:01	36.16	21.85
AATAQTLQAFLHWAI	P15712.1	HLA-DRB1*15:01, HLA-DRB1*07:01	34.78	9.37
EEYLILEDKILVQAN	P09621.3	HLA-DRB1*03:01, HLA-DPA1*03:01/DPB1*04:02, HLA-DRB1*04:01, HLA-DRB4*01:01, HLA-DPA1*02:01/DPB1*14:01, HLA-DRB5*01:01	27.97	13.51
AGVQYSRADVVRFQE	NP_218391.1	HLA-DRB3*02:02, HLA-DRB1*13:02, HLA-DRB1*07:01	24.27	21.15
AAVVRFQEAANKQKQ	NP_218391.1	HLA-DRB5*01:01, HLA-DRB1*08:02, HLA-DRB1*04:01, HLA-DRB1*11:01,	23.19	26.73
DQVHFQPLPPAVVKL	P15712.1	HLA-DRB1*01:01, HLA-DRB1*09:01, HLA-DRB1*12:01	21.60	16.64
AAWGGWNFAGIEAAA	P0A564.2	HLA-DRB1*11:01, HLA-DRB1*04:05, HLA-DRB1*09:01	19.34	32.92

Formatted: Font: 11 pt
Formatted: Font: 11 pt
Formatted: Font: 11 pt
Formatted: Font: 11 pt
Formatted: Font: 11 pt
Formatted: Font: 11 pt
Formatted: Font: 11 pt
Formatted: Font: 11 pt
Formatted: Font: 11 pt
Formatted: Font: 11 pt

AAVLTDASATKDGSH	P0A5J0.1	HLA-DRB1*11:01, HLA-DRB3*01:01	10.54	26.73
CVAYIGISFLDQASQ	P15712.1	HLA-DPA1*01:03/DPB1*02:01, HLA-DQA1*01:01/DQB1*05:01, HLA-DPA1*01/DPB1*04:01, HLA-DPA1*02:01/DPB1*01:01, HLA-DPA1*02:01/DPB1*14:01, HLA-DQA1*05:01/DQB1*02:01, HLA-DRB1*11:01, HLA-DRB4*01:01, HLA-DPA1*03:01/DPB1*04:02	10.54	26.73

Formatted: Font: 11 pt

Formatted: Font: 11 pt

TABLE 3: Characteristics of Epitope combinations

The optimal combinations of conserved CD8+ and CD4+ epitopes are listed below. This Table lists the sequences of the conserved CD8+ and CD4+ epitopes ranked on their overall Projected Population Coverage or PPC, given as a percentage, for both the world and for East Africa. The PPC is calculated from the prevalence of alleles predicted as binding to the given epitope, see tables 1 and 2.

EPITOPE COMBINATION	PPC (%)	
	WORLD	EAST AFRICA
QSSFYSDW, ALAALGLWLSV, MLLAVTVSL, AEQEQCLDEL, FPAGGSTGSL, HISSGVFLLK, KMRCGAPRY, WYYQSGLSI, NTPAFEWYY	97.76	86.14
SRGWSLIKSVRLGNA, KPRIITLTMNPALDI, AAHKGLMNIALAISA, FPAGGSTGSL, MLLAVTVSL, QSSFYSDW, KMRCGAPRY	97.39	92.67
QSSFYSDW, MLLAVTVSL, AEQEQCLDEL, FPAGGSTGSL, HISSGVFLLK, KMRCGAPRY, WYYQSGLSI	97.11	85.63
SRGWSLIKSVRLGNA, KPRIITLTMNPALDI, AAHKGLMNIALAISA, FPAGGSTGSL, MLLAVTVSL, QSSFYSDW	96.09	88.27
SRGWSLIKSVRLGNA, KPRIITLTMNPALDI, AAHKGLMNIALAISA, FPAGGSTGSL, MLLAVTVSL	93.72	83.80